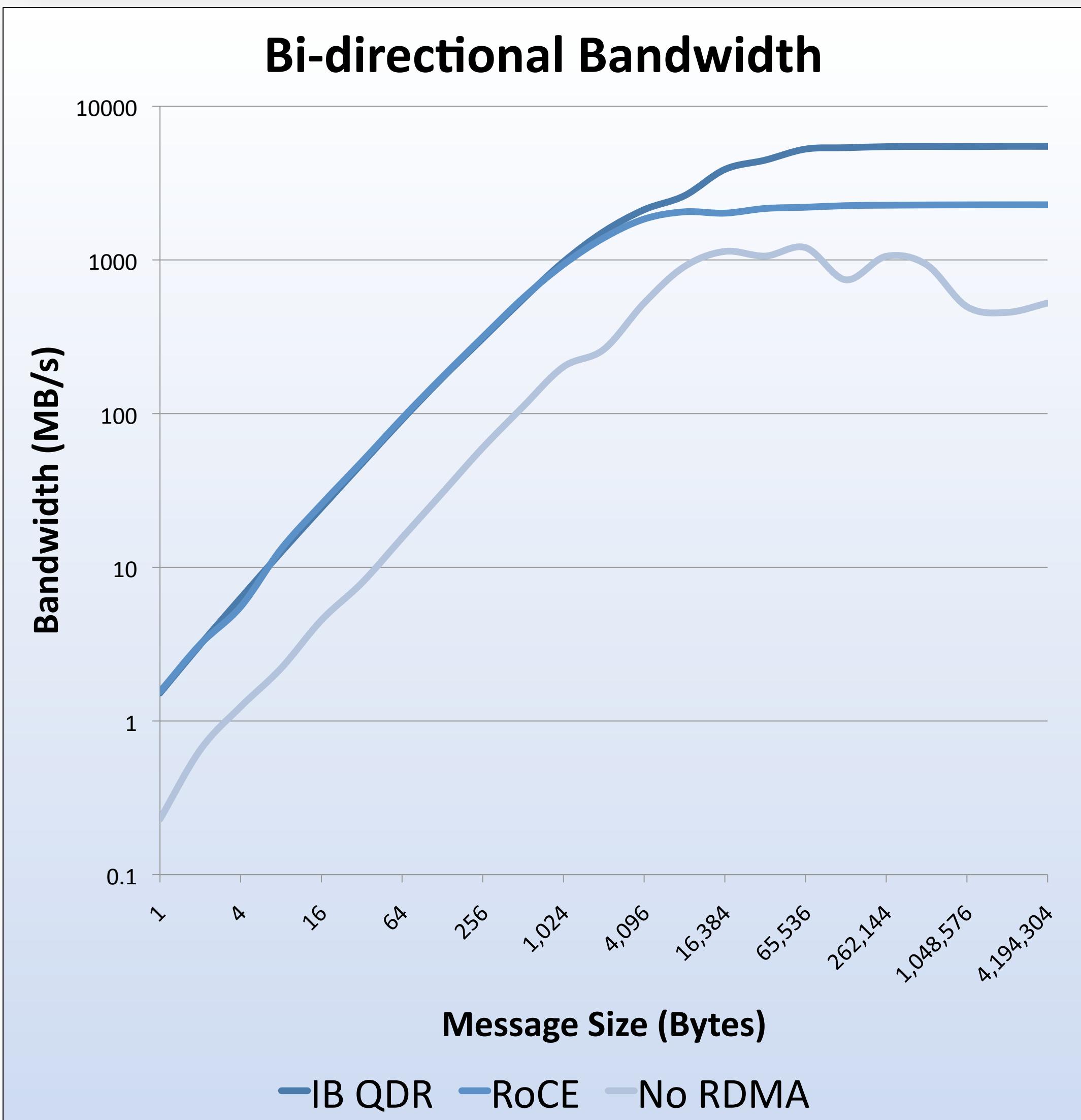
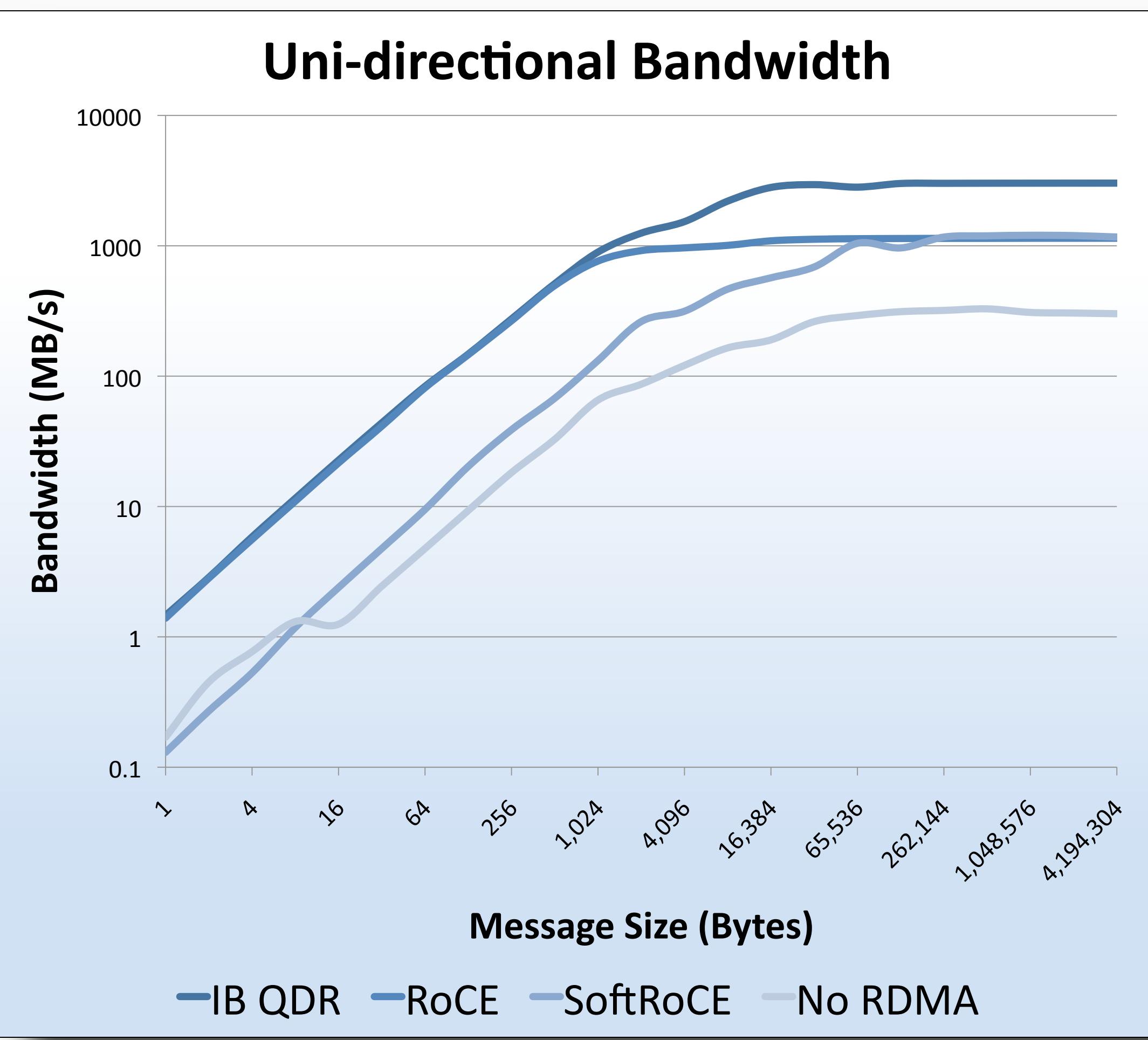


Implementation & Comparison of RDMA Over Ethernet



Student Researchers

Lee Gaiser
lgaiser@mtu.edu

Brian Kraus
bkraus@unm.edu

James Wernicke
jwernick@nmt.edu

Mentors

Andree Jacobson
andree@lanl.gov

Susan Coulter
skc@lanl.gov

Jharrod LaFon
jlafon@lanl.gov

Ben McClelland
ben@lanl.gov

Special Thanks

Samuel Gutierrez
samuel@lanl.gov

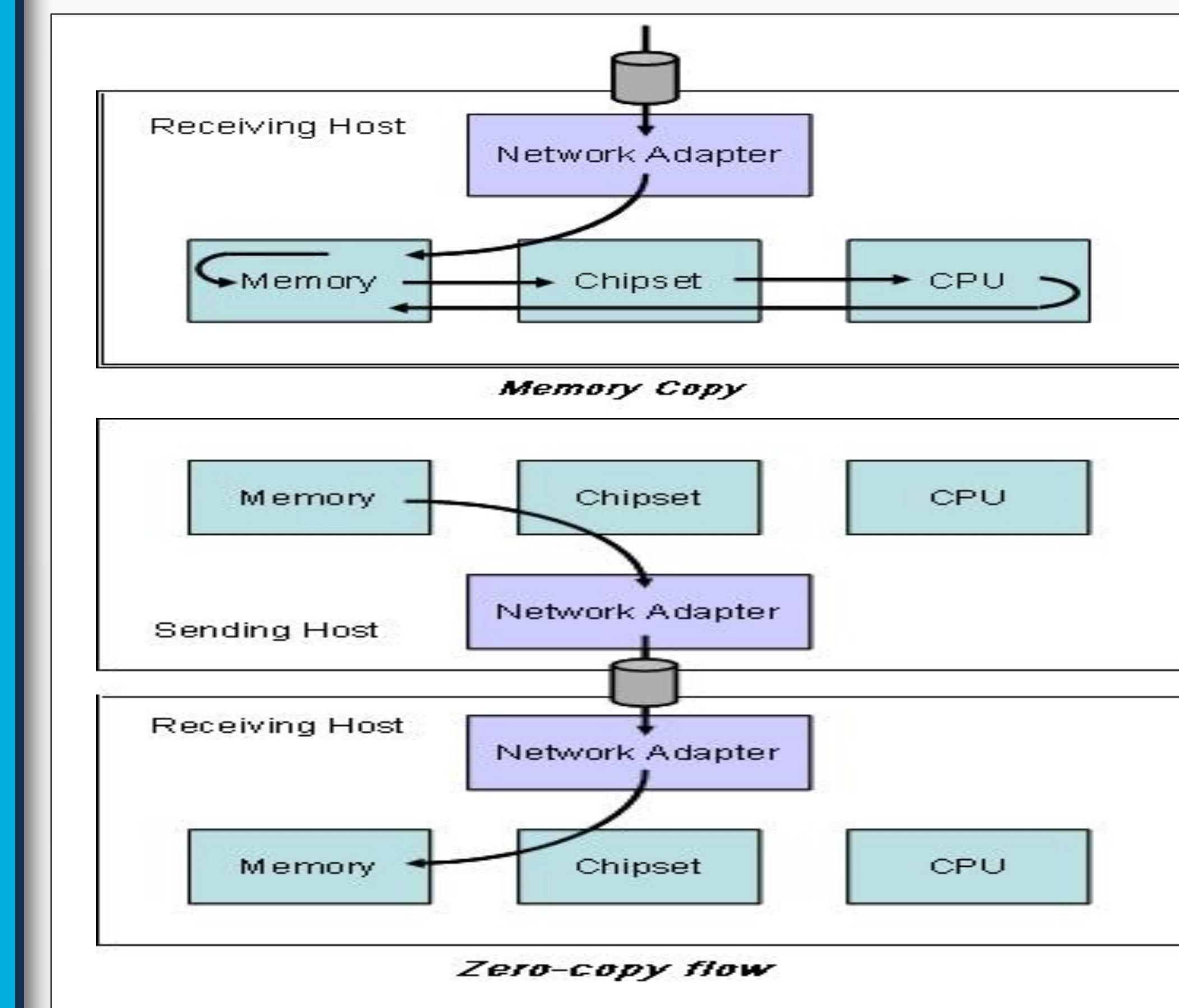
Robert Pearson
rpearson@systemfabricworks.com



Abstract

Remote direct memory access (RDMA) allows a computer's network adapter to transfer data directly to and from its application memory. This process significantly improves latency in cluster environments by eliminating the transfer of data between application memory and kernel memory as well as reducing network protocol overhead.

Because many existing networks were built with Ethernet, RDMA over Converged Ethernet (RoCE) can significantly increase bandwidth without the cost of upgrading infrastructure. Additionally, software-only implementations of RoCE can eliminate the cost of installing new NICs. In this project, we implement RoCE both in software and hardware, quantify the performance, and compare that performance to non-RDMA Ethernet and Infiniband RDMA. Analysis of the hardware being used and the cost & benefit of implementation choices are also part of this project.



Images courtesy of <http://www.hpcwire.com/features/17888274.html>

Results & Analysis

Peak Values	IB QDR	RoCE	Soft RoCE	No RDMA
Latency (μs)	1.96	3.7	11.6	21.09
One-way BW (MB/s)	3024.8	1142.7	1204.1	301.31
Two-way BW (MB/s)	5481.9	2284.7	-	1136.1

➤ RoCE performance gains over non-RDMA 10GbE:

- Up to 5.7x speedup in latency
- Up to 3.7x increase in bandwidth

➤ RoCE vs. IB QDR:

- Less than 1μs slower at 128-byte messages
- Comparable bandwidth up to 1KB messages
- IB peak bandwidth 2-2.5x greater than RoCE at very large messages

